Assessing Local Variations of Deforestation Processes in Mexico Using Geographically Weighted Regression

Jean-François Mas*, Gabriela Cuevas*, Araceli Andablo Reyes* **, Frida Guïza*

* Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México (UNAM)
** Centro de Investigación en Alimentación y Desarrollo A.C. - CIAD

Abstract. This study identifies drivers of deforestation in Mexico by applying Geographically Weighted Regression (GWR) models to cartographic and statistical data. A wall-to-wall multitemporal GIS database was constructed incorporating digital land use/land cover maps for 2002 and 2007; along with ancillary data (road network, settlements, topography and socio-economical parameters). The database analysis allowed assessing the spatial distribution of forests and deforestation at the municipal level. The statistical analysis of deforestation drivers presented here was focused on the proportion of anthropogenic cover in 2007 as dependent variable. In comparison with the global model, the use of GWR increased the strength in the relationship in terms of the goodness-of-fit (adjusted R^2) from 0.69 (global model) to 0.72 (average R^2 of GWR local models), with individual GWR models ranging from 0.48 to 0.81. The GWR model highlighted the spatial variation of the relationship between the percentage of anthropogenic cover and its drivers. Factors identified as having a major impact on deforestation were related to topography (slope), accessibility (road and settlement density) and marginalization. Results indicate that some of the drivers explaining deforestation vary over space, and that the same driver can exhibit opposite effects depending on the region. Based on local regression model coefficients, a cluster analysis allowed the aggregation of municipalities with similar patterns of deforestation into homogeneous regions. A deforestation model for the entire country will be developed further, using these regions to divide the model procedures into sub-regions with specific deforestation patterns.

Keywords: Land Use/Cover Change, Geographically Weighted Regression, Mexico

1. Introduction

Mexico, with a total area a total area of about 2 million square kilometers, is a megadiverse country, but it presents high rates of deforestation (FAO 2001). Various studies have attempted to assess land use / cover change (LUCC) over the last decades (Mas et al. 2004) but there have been few attempts to assess the main causes of deforestation at national level (Figueroa et al. 2009, Pineda Jaimes et al 2010, Bonilla-Moheno et al. 2013). Geographically Weighted Regression (GWR) have been applied in exploring spatial data in the social, health and environmental sciences. The goal of this study is to evaluate the spatial patterns of deforestation with respect to drivers reported to influence LUCC using Geographically Weighted Regression (GWR). In this paper, we present the preliminary results.

2. Material and Methods

2.1. Material

In order to elaborate the GIS database, the following data were used:

• Maps of land use/cover (LUC) at 1:250,000 scale from the National Institute of Geography, Statistics and Informatics (INEGI for its Spanish acronym) for 2002 and 2007. These maps are compatible with regards to scale and classification scheme (Mas et al. 2004).

• Maps of ancillary data (digital elevation model, roads maps, human settlements, municipal boundaries).

• Socio-economic data from the INEGI organized by municipality (Population census for 2000, 2005 and 2010).

GIS operations were carried out with the following programs: ArcGIS (ESRI, Redlands, CA) and Q-GIS (www.qgis.org/). Statistical analysis and graphs were created using R (R Development Core Team 2009). Geographically weighted regressions (GWR) were carried out using the packages: gwrr (Wheeler 2007 and 2012) and spgwr (Bivand and Yu 2012) in R.

2.2. LUCC Monitoring and GIS database elaboration

LUCC monitoring was done by overlaying the LUC maps of 2002 and 2007 and the map of municipalities. Areas of change were tabulated and computed as well as the rate of deforestation. In this study, the rate of deforestation reflects the transformation of wooded covers (temperate and tropical forests, scrublands) into anthropogenic covers (agriculture, pasture land, urban areas). Two dependent variables were generated for each municipality: the proportion of area covered by anthropogenic covers in 2007 as a proxy

for the level of anthropization and the rate of deforestation between 2004 and 2007. The rate of deforestation were obtained from municipalities with a forested area covering at least 500 ha and 10% of the municipality in order to avoid outlier values. More than 2106 municipalities out of 2456 fulfilled these conditions, representing more than 97.2 % of the country total area and 99.8% of the forest and scrubland area.

In order to determine which ancillary variables are most likely to be indirect drivers of deforestation, we calculated, for each municipality various indices describing: population, land tenure, economic activities and the resources accessibility. These indices were: a) Population density in 2010 (people per km^2); b) Density settlements (number of settlements per km^2); c) Index of marginalization (CONAPO 2010), d) Cattle density, e) Goat density, f) Mean slope (degrees), g) Road density (km of road per km^2), h) Proportion of municipality area with private tenure, i) Proportion of area with *ejidal* tenure, and j) Proportion of area with *communal* tenure. *Ejidal* and *Communal* tenures are both common-pool systems, related to land redistribution process and land title restitution respectively.

2.3. Statistical Analysis

Geographically Weighted Regression is a local spatial statistical technique for exploring spatial nonstationarity (Fotheringham et al., 2002). It supports locally modeling of spatial relationships by fitting regression models. Regression parameters are estimated using a weighting function based on distance in order to assign larger weights to closer locations. Different from the usual global regression, which produces a single regression equation by summarizing the overall relationships among the explanatory and dependent variables (for the whole Mexican territory in that case), GWR produces spatial data that express the spatial variation in the relationships among variables. Maps that present the spatial distribution of the regression coefficient estimates along with the level of significance (e.g. t-values) have an essential role in exploring and interpreting spatial nonstationarity. Fotheringham et al. (2002) provide with a full description of GWR, and Mennis (2006) gives useful suggestions to map GWR results.

The first stage of the study was correlation analysis between explanatory variables using the Spearman coefficient in order to discard highly correlated variables. Due to the uneven distribution and size of the municipalities, the weighting function was based on a proportion of the observations (k-nearest neighbors) assigned to the centroid of each municipality. The optimal size of the bandwidth was evaluated by minimizing the root mean square error. A map was elaborated for each explanatory variable showing the value of the regression's coefficients (color scaling of the centroid sym-

bol) and statistical significance (size of the centroid symbol). Next, a cluster analysis was carried out assorting municipalities with common deforestation patterns (similar regression coefficients) into regions. Finally, a global regression model was applied to each region.

3. Results

3.1. LUCC Monitoring

According to INEGI maps, anthropogenic covers (mainly crop and pasture lands) occupy about 26% of the Mexican territory. During 2002-2007, the anthropogenic cover area increased (14,600 km^2), while tropical forest and scrubland area decreased (reduction of 8400 and 5200 km^2 respectively). The area of temperate forest remains broadly the same. As shown in Figure 1 the Anthropization index varies over space: In the coastal floodplains of the Gulf of Mexico and the center of Mexico, original vegetation cover has been almost totally removed while other regions are better conserved.

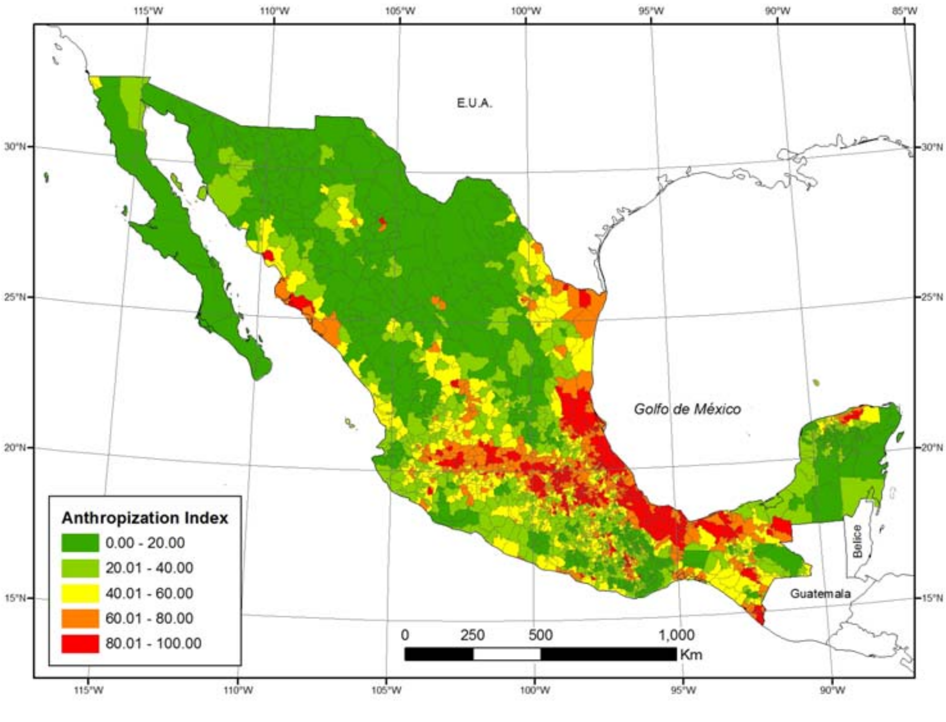

Figure 1. *Per municipalty* Anthropization index (2007)

3.2. Geographically Weighted Regression (GWR)

In this paper, we report only the results of the GWR using as dependent variable the proportion of anthropogenic cover (anthropization index). The weighting function was based on a 5% of the observations. A global model was fitted and obtained an adjusted-R^2 of 0.69. The use of GWR slightly increased the strength in the relationship in terms of the goodness-of-fit (adjusted R^2) to 0.72 (average R^2 of GWR local models), with local GWR models with adjusted R^2 ranging from 0.48 to 0.81. Figure 2 presents the spatial distribution of the goodness of fit.

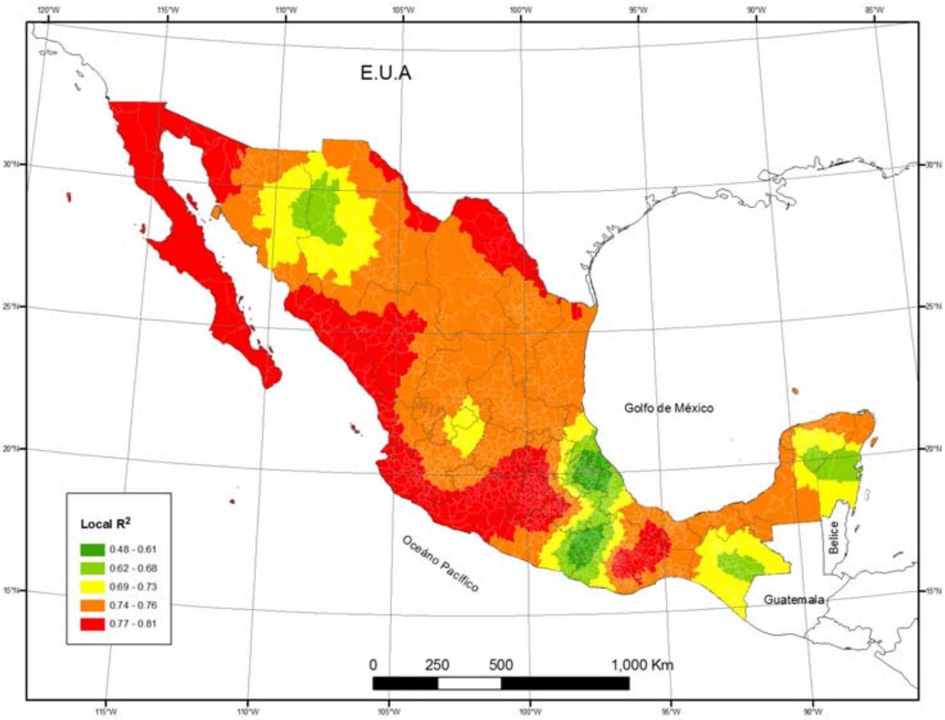

Figure 2. Distribution of local R^2.

Some variables such as population and road density and slope exhibit a significant relationship with the proportion of anthropogenic cover for the whole territory. As expected, the first two variables have a positive effect on this proportion while slope presents a negative relationship (Figure 3). Other explanatory variables have a more contrasted patterns. In example, the marginalization index presents a significant relationship with the proportion of anthropogenic cover only for about half of the territory. It presents a positive relationship in the Baja California states, the border with the USA

and the north strip along the Gulf of Mexico and a negative relationship in various regions of Mexico Southeastern part (Figure 3).

The region where the relationship is positive is related to industrialized regions where more developed municipalities present both higher proportion of anthropogenic cover and lower marginalization index. Opposite to that, in the southeastern region of the country more conserved areas are associated to higher levels of marginalization. Previous researches have reported that the most conserved natural areas in Mexico are often located in poor rural areas and/or community lands (Klooster 2000; Alix-Garcia et al. 2005, Figueroa et al. 2009, García-Barrios et al. 2009).

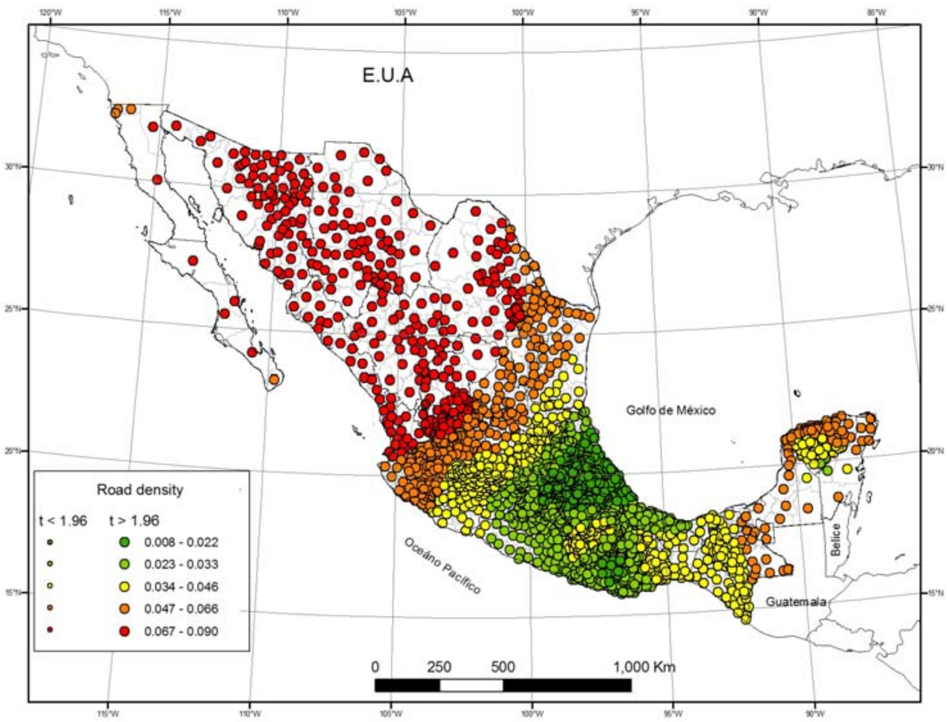

Figure 3. GWR coefficient and significance values for road density.

The cluster analysis helped dividing the whole territory into five principal regions (Figure 5). It is worth noting that some of these regions have a high coincidence with other territory zoning such as physiographic map (INEGI 1981). Table 1 presents the variables retained in each regional regression model, along with the sign and the significance level of their coefficient.

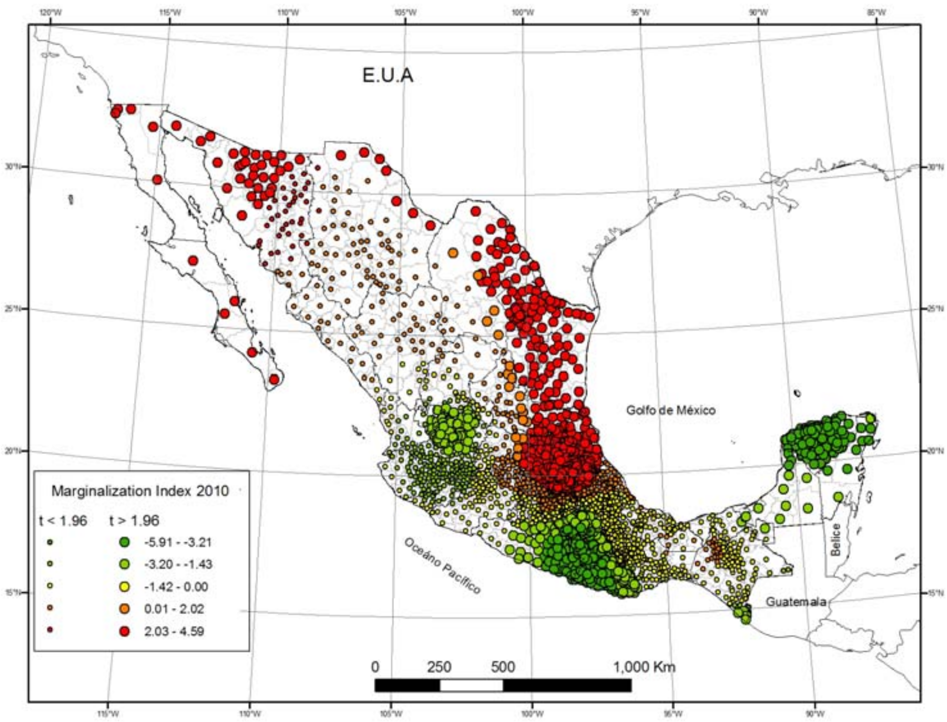

Figure 4. GWR coefficient and significance values for marginalization.

Region Number	1	2	3	4	5
Marginalization index		- ***	- ***		+ **
Settlements density	+ ***	+ ***	+ ***	+ ***	+ ***
Population density	+ **				+ .
Cattle density				+ ***	+ ***
Goat density	+ **	+ ***	- ***	+ **	+ **
Roads density	+ ***	+ ***	+ ***	+ ***	+ ***
NPA proportion	- ***			- *	- *
Wood production	+ .		+ .	+ *	+ *
Slope	- ***	- ***	- ***	- ***	- ***
Land tenure *Ejidal*		- *		- ***	
Land tenure *Communal*			- **	- ***	- ***
Land tenure Private			- ***	- ***	- *

Table 1. Sign and significance of coefficients region-based regression. Significance codes: 0 ***, 0.001 **, 0.01 *, 0.05 .

Figure 5. Regions based on the cluster analysis of GWR coefficients.

4. Discussion and Conclusion

Some limitations of this study have been identified: First, it is based only on a drastic change of land cover (forested cover *versus* anthropogenic cover), it does not consider cover degradation. This factor has to be considered during the results interpretation. For example in some regions goat density is associated with lower levels of anthropization, however it is likely related with vegetation cover degradation considered in this study as natural (without any indication about degradation level). Second, we used contemporary explanatory variables (e.g. 2005 and 2010 population census) to explain anthropization patterns resulting from centuries of landscape transformation. In fact, it is likely that the effect of a driver on a given region is related to the time such driver has been shaping the landscape. Finally, as depicted in figure 2, the set of explanatory variables we used did not allow to explain the dependent variable in a satisfactory manner for the entire territory. More drivers have to be taken into account for future analysis.

However, results clearly show the advantages of a local approach (GWR) over a global one, to assess different drivers' effect on LUCC over such a complex and diverse territory as Mexico. In future researches, explanatory variables will be integrated into the model to test the effect of migration and public policies. Considering the rate of deforestation during different past periods of time (instead of the level of anthropization) will enable us to analyze the dynamic of deforestation in its temporal and spatial dimensions. A deforestation model for the entire country will be developed, using the suggested regions (figure 5) dividing the model procedures into sub-regions with specific deforestation patterns. Spitting the model into sub-regions is expected to improve the prospective map accuracy produced by the model.

Acknowledgments

This research has been funded by the *Consejo Nacional de Ciencia y Tecnología* (CONACyT) and the *Secretaría de Educación Pública* (grant CONACYT-SEP CB-2012-01-178816) and *Dirección General de Asuntos del Personal Académico* (DGAPA) at the *Universidad Nacional Autónoma de México* (UNAM) (grant RR113511). Land use/cover map, scale 250,000 was provided by the INEGI in the framework of the agreement between UNAM and INEGI (number DGAD-036/2011).

References

Alix-Garcia J, de Janvry A, Sadoulet E (2005) A Tale of Two Communities: Explaining Deforestation in Mexico. World Development, 33(2): 219-235.

Bivand R, Yu D (2012) Package 'spgwr', Geographically weighted regression. http://cran.open-source-solution.org/web/packages/spgwr/spgwr.pdf Accessed 12 April 2013

Bonilla-Moheno M, Redo DJ, Mitchell Aide T, Clark ML, Grau HR (2013) Vegetation change and land tenure in Mexico: A country-wide analysis. Land Use Policy 30(1): 355-364.

CONAPO (2010) Índices de marginación por localidad. http:// http://www.conapo.gob.mx/es/CONAPO/Indice_de_Marginacion_por_Localidad_2010 Accessed 15 April 2013

FAO (2001) Global resources assessment. Forestry paper 140.

Figueroa F, Sánchez-Cordero V, Meave JA, Trejo I (2009). Socioeconomic context of land use and land cover change in Mexican biosphere reserves. Environmental Conservation 36 (3): 180–191.

Fotheringham SA, Brunsdon C, Charlton M (2002) Geographically Weighted Regression: the analysis of spatially varying relationships. John Wiley & Sons.

García-Barrios L, Galván-Miyoshi YM, Valdivieso Pérez IA, Masera OR, Bocco G, Vandermeer J (2009) Neotropical forest conservation, agricultural intensification and rural outmigration: The Mexican Experience. BioScience, 59(10): 863-873.

INEGI (1981) Carta fisiográfica de la República Mexicana, escala 1:1000000.

Klooster D (2000) Beyond Deforestation: The Social Context of Forest Change in Two Indigenous Communities in Highland Mexico. Journal of Latin American Geography 26: 47-59.

Mas JF, Velázquez A, Díaz-Gallegos JR, Mayorga-Saucedo R, Alcántara C, Bocco G, Castro R, Fernández T, Pérez-Vega A (2004) Assessing land/use cover changes: a nationwide multidate spatial database for Mexico. International Journal of Applied Earth Observation Geoinformatics 5: 249–261.

Mennis JL (2006) Mapping the results of geographically weighted regression. The Cartographic Journal 43(2): 171-179.

Pineda Jaimes NB, Bosque Sendra J, Gómez Delgado M, Franco Plata R (2010) Exploring the driving forces behind deforestation in the state of Mexico(Mexico) using geographically weighted regression. Applied Geography 30: 576–591.

R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. http://www.R-project.org Accessed 9 April 2013

Wheeler DC (2007) Diagnostic tools and a remedial method for collinearity in geographically weighted regression. Environment and Planning A, 39: 2464-2481.

Wheeler DC (2012) Package ‘gwrr’, Geographically weighted regression with penalties and diagnostic tools Austria. http://cran.r-project.org/web/packages/gwrr/gwrr.pdf Accessed 9 April 2013